FEBS 28349 FEBS Letters 566 (2004) 17–24

Characterization of a novel human UDP-GalNAc transferase, pp-GalNAc-T15 ☆

Lamei Cheng[a,1], Kouichi Tachibana[a], Hiroko Iwasaki[a,b], Akihiko Kameyama[a,b], Yan Zhang[a], Tomomi Kubota[a], Toru Hiruma[a,c], Kahori Tachibana[a], Takashi Kudo[a,2], Jian-Ming Guo[a,3], Hisashi Narimatsu[a,*]

[a] *Glycogene Function Team, Research Center for Glycoscience, National Institute of Advanced Industrial Science and Technology (AIST), Central-2, 1-1-1 Umezono, Tsukuba, Ibaraki 305-8568, Japan*
[b] *Amersham Biosciences KK, 3-25-1, Hyakunincho, Shinjuku-ku, Tokyo 169-0073, Japan*
[c] *Fujirebio Inc., Diagnostics Reserch Laboratories, 51 Komiyacho, Hachioji, Tokyo 192-0031, Japan*

Received 27 January 2004; revised 23 March 2004; accepted 29 March 2004

Available online 22 April 2004

Edited by Masayuki Miyasaka

Abstract **We have cloned, expressed and characterized a novel member of the human UDP-GalNAc:polypeptide *N*-acetylgalactosaminyltransferase (pp-GalNAc-T) family, pp-GalNAc-T15. The *pp-GalNAc-T15* transcript was ubiquitously expressed in human tissues. Recombinant pp-GalNAc-T15 transferred *N*-acetylgalactosamine (GalNAc) toward a panel of mucin-derived peptide substrates in vitro. Although pp-GalNAc-T15 showed significantly less catalytic activity than pp-GalNAc-T2, T15 transferred up to seven GalNAcs to the Muc5AC peptide, while T2 transferred up to five GalNAcs. These results clearly indicated that pp-GalNAc-T15 is a novel member of the human pp-GalNAc-T family with unique catalytic activity.**

Keywords: UDP-GalNAc:polypeptide *N*-acetylgalactosaminyltransferase (EC 2.4.1.41); O-glycosylation; *O*-Glycan; Mucin

1. Introduction

Mucin-type O-linked glycosylation is initiated by the addition of *N*-acetylgalactosamine (GalNAc) to serine or threonine residues of the polypepide chain. This reaction is known to be catalyzed by a family of UDP-GalNAc:polypeptide *N*-acetylgalactosaminyltransferases (pp-GalNAc-Ts) (EC 2.4.1.41) [1], which transfer GalNAc from UDP-GalNAc to a serine or threonine residue on the polypeptide acceptor. To date, 14 distinct mammalian isoforms of this gene family (T1–T14) [1–17] have been identified and functionally characterized. These isoforms have been shown to have unique acceptor substrate specificities as well as tissue-specific expression patterns; several isoforms are expressed in various tissues and catalyze a broad spectrum of substrates (pp-GalNAc-T1, T2), whereas the other isoforms are more restricted in expression and/or in substrate preference (pp-GalNAc-T3, T4, T7, T9, T11 and T13). The differences in kinetic properties, substrate specificities, and expression patterns of pp-GalNAc-Ts are presumably involved in a hierarchical order of the mucin-type O-glycosylation processing and govern *O*-glycan attachment sites and density on peptide backbones in vivo. Although it is still hard to define the in vivo catalytic specificity and biological function of an individual GalNAc-T isoform, a few in vivo studies have confirmed the in vitro characterization of catalytic activity, mainly characteristic sequence motifs of acceptor substrates [18]. Thus, to understand how O-glycosylation is initiated and regulated in a given tissue or cell type, the repertoire and catalytic specificity of pp-GalNAc-Ts in that cell background have to be defined.

In the present study, we report the cloning and characterization of pp-GalNAc-T15, a novel member of this GalNAc-transferase family. For the characterization, we compared the catalytic specificity of pp-GalNAc-T15 with that of T2, and found it to be totally different.

2. Materials and methods

2.1. Cloning of pp-GlaNAc-T15 cDNA

A BLAST search of expressed sequence tag (EST) databases identified three cDNAs (GenBank™ Accession Nos.: BG699346, BG714178 and BG715977) as homologues to the open reading frame

☆ The nucleotide sequence of human pp-GalNAc-T15 reported in this paper has been deposited in the DDBJ/EMBL/GenBank databases under Accession No. AB078149.

* Corresponding author. Fax: +81-298-61-3201.
E-mail address: h.narimatsu@aist.go.jp (H. Narimatsu).

[1] Present address: Experimental Hematology Laboratory, Xiang Ya Medical School, Central South University, Changsha 410078, PR China.
[2] Present address: Department of Anatomy and Embryology, Institute of Basic Medical Sciences, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8575, Japan.
[3] Present address: Department of Urology, Zhongshan Hospital, Medical Center, Fu-Dan University, 136 Yi Xue Yuan Road, Shanghai 200032, PR China.

Abbreviations: pp-GalNAc-T, UDP-GalNAc:polypeptide *N*-acetylgalactosaminyltransferases; ORF, open reading frame; PCR, polymerase chain reaction; HPLC, high-performance liquid chromatography; FAM, 5-carboxyfluorescein succinimidyl ester; TFA, trifluoroacetic acid; MALDI-TOF, matrix-assisted laser desorption/ionization time-of-flight; PTH, phenylthiohydantoin

0014-5793/$22.00
doi:10.1016/j.febslet.2004.03.108

Table 1
Acceptor peptide specificities of purified recombinant pp-GalNAc-T15 and -T2

Acceptor peptide	Amino acid sequence	GalNAc-T15 (U/μg)	GalNAc-T2 (U/μg)
Muc1a	AHGVTSAPDTR	0.023	0.24
Muc2	PTTTPITTTTTVTPTPTPTPTGTQT	0.008	7.94
Muc5AC	GTTPSPVPTTSTTSAk	0.0126	8.37
Muc13	ANTPSFPTATSPAPPI	0.012	2.98
EA2	PTTDSTTPAPTT	0.006	3.06

Relative activity determined in reactions containing 200 μM peptide substrate and purified GalNAc-T15 and -T2.

(ORF) of human pp-GalNAc-T7 (AJ002744). By searching the human genomic DNA database, we found a genomic sequence (NT005927) that contained a putative initial codon and a transmembrane domain of this putative gene. A full-length ORF of a novel polypeptide GalNAc transferase was identified by database gene walking and deposited into GenBank with the Accession No. AB078149. The cDNA encoding the full-length ORF was obtained by polymerase chain reaction (PCR) using the Expand™ High Fidelity PCR system (Roche), the Marathon-ready™ cDNA of human cerebellum (Clontech) as a template, and the primers, 5′-GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGCTCCTAAGGAAGCGATACAGGCAC-3′ (forward) and 5′-GGGGACCACTTTGTACAAGAAAGCTGGGTCCTCATCGTTCATCCACAGCATTGATCT-3′ (reverse). These primers have an attB1 or an attB2 overhang to create recombination sites. The amplified fragment was subcloned into the vector pDONR201™ using the BP CLONASE Enzyme Mix (Invitrogen). The sequences of subcloned cDNA were determined using a DYEnamic™ DNA Sequencing Kit (Amersham Pharmacia Biotech).

2.2. Quantitative analysis of the pp-GalNAc-T15 transcripts in human tissues by real-time PCR

Real-time PCR was carried out using the TaqMan Universal PCR Master Mix (Applied Biosystems), Marathon-Ready cDNAs of various human tissues (Clontech), the forward primer (5′-GGATGTGCTCGTCTTCATGGA-3′), the reverse primer (5′-GATACCACTCGGCTCCTGTCA-3′), and the probe (5′-CTGGAGCCCCTCCTCAGCAGAATAGC-3′). Standard curves were generated by methods described previously [19]. PCR products were continuously measured with the ABI PRISM 7700 Sequence Detection System (Applied Biosystems). The relative amount of *pp-GalNAc-T15* transcript was normalized to the amount of human *glyceraldehyde-3-phosphate dehydrogenase* transcript in each cDNA.

2.3. Production and purification of recombinant pp-GalNAc-T15 protein with FLAG peptide tag

The cDNA fragment encoding the putative catalytic domain of pp-GalNAc-T15 (amino acids 45–639) was amplified by PCR using the primers, 5′-GGGGACAAGTTTGTACAAAAAAGCAGGCTTCGCCAGGTACCGCCTGGACTTTGGGG-3′ (forward) and 5′-GGGGACCACTTTGTACAAGAAAGCTGGGTCCTCATCGTTCATCCACAGCATTGATCT-3′ (reverse). The amplified cDNA fragment was cloned into pDONR201™ and subcloned into the expression vector pFBIF using the LR CLONASE Enzyme Mix (Invitrogen) [13,19]. The pFBIF–pp-GalNAc-T15 truncated construct was tranformed into $DH10_{BAC}$ competent cells (Invitrogen) to generate bacmid DNA. The bacmid was introduced into Sf21 cells (Pharmingen) to yield conditioned medium containing recombinant pp-GalNAc-T15 protein. The secreted recombinant protein was purified using anti-FLAG M1 antibody-conjugated resin (Sigma). The purified protein was fractionated by 10% SDS–PAGE, and analyzed by Western blotting using anti-FLAG M2 (Sigma) and the enhanced chemiluminescence-plus (ECL^+) Western Blot Detection Kit (Amersham Biosciences). The relative concentration of the purified protein was estimated using the FLAG-tagged bacterial alkaline phosphatase (Sigma) as a standard.

The soluble form of human pp-GalNAc-T2, containing amino acid residues 52–571 [2], was also generated, purified and quantified by the same methods as for pp-GalNAc-T15.

2.4. Polypeptide GalNAc-transferase assays

The standard enzyme reaction mixture (final volume, 20 μl) consisted of 25 mM Tris–HCl (pH 7.4), 10 mM $MnCl_2$, 0.2% Triton X-100, 1 mM UDP-GalNAc (Sigma), 5 mM 2-mercaptoethanol, 0.005–2 mM acceptor peptides labeled with 5-carboxyfluorescein succinimidyl ester (FAM), and an adequate amount of purified pp-GalNAc-Ts. Reactions were performed at 37 °C for various periods and were terminated by boiling. FAM-labeled peptide substrates: Muc1a, Muc2, Muc5AC, Muc13 and EA2 [20–24] (see Table 1 for structure), which were derived from the tandem repeats of corresponding mucins, were purchased from Sawady Co. Ltd (Tokyo, Japan). The reaction products were evaluated by high-performance liquid chromatography (HPLC) on a C_{18} reverse column (Water $5C_{18}$-AR, 4.6×250 mm) as previously described [25].

For the initial velocity assay of pp-GalNAc-Ts with a variety of substrate peptides, the amount of substrate consumed was limited to less than 10% of the total. The relative activity of pp-GalNAc-Ts with each substrate peptide was presented as units per micrograms. One unit of enzyme is defined as the amount of enzyme that transfers 1 nmol of GalNAc in 1 min using the standard reaction mixture. The GalNAc-transferase assay, to determine the K_m for the acceptor substrate Muc1a, was modified to include 1 mM UDP-GalNAc with Muc1a at concentrations from 0.005 to 2 mM. Assays to determine the K_m for UDP-GalNAc were performed with the saturating concentration of Muc1a peptide (pp-GalNAc-T15, 1 mM Muc1a; pp-GalNAc-T2, 500 μM Muc1a). Assays were performed in duplicate or quadruplicate.

2.5. Separation of enzyme reaction products

The glycosylated peptides were isolated by HPLC. As a standard separation, the reaction products were eluted with 13% solvent B (0.05% trifluoroacetic acid (TFA) in 70% 2-propanol in acetonitrile) in solvent A (0.05% TFA in water) at a flow rate of 1 ml/min for 60 min. Eluates were monitored by the fluorescence intensity at 520 nm (excitation: 492 nm).

2.6. Identification of reaction products by matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) spectrometry and peptide sequencing

The isolated fractions were dried, redissolved in distilled water, mixed with the same amount of matrix solution (α-cyano-4-hydroxycinnamic acid, saturated in 30% acetonitrile containing 0.1% TFA), and subjected to MALDI-TOF mass spectrometry using REFLEX™ IV (Bruker Daltonics) in the negative ion mode. The method was described in our previous report [25].

To identify the positions of GalNAc-attached residues, amino acid sequencing of glycopeptides by Edman degradation was performed with a peptide sequencer (PPSQ-23A, Shimadzu). With this sequencing system, a phenylthiohydantoin (PTH) derivative of GalNAc-attached Thr (Calbiochem, Darmstdt, Germany) was identified as a pair of peaks eluted at 3.07 and 3.25 min, which were close to the DTT peak, and that of GalNAc-attached Ser was identified as a peak at 2.96 min, which overlapped the PTH-Glu peak.

3. Results

3.1. cDNA cloning of pp-GalNAc-T15

The *pp-GalNAc-T15* gene was identified as a putative novel *pp-GalNAc-T* gene by searching the EST database with the *human pp-GalNAc-T7* gene as a query, and the cDNA was cloned as described in Section 2. The nucleotide and deduced amino acid sequences of pp-GalNAc-T15 are shown in Fig. 1A, and an alignment of the amino acid sequences of 13

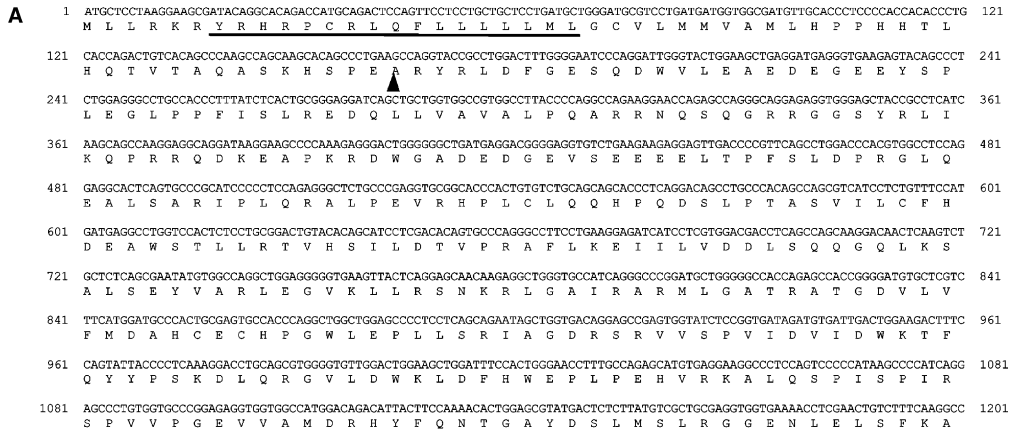

Fig. 1. (A) Nucleotide and predicted amino acid sequences of pp-GalNAc-T15. Numbering of the cDNA begins with the initiation codon. The predicted amino acid sequence is presented under the putative nucleotide coding sequence. The *triangle* represents the starting position of the truncated form. The predicted transmembrane domain is underlined with a *single solid line*. (B) ClustalW alignment for comparison of pp-GalNAc-T15 with the other pp-GalNAc-Ts. The putative transmembrane domains are underlined. Three motifs, GT1, Gal/GalNAc transferase and lectin-like, are indicated above the sequences. The amino acids that are identical among all proteins are shown by asterisks, while conserved amino acids are shown by dots.

pp-GalNAc-Ts generated using ClustalW is shown in Fig. 1B. The cDNA encoding pp-GalNAc-T15 contains an ORF of 1917 bp encoding a 639-amino acid type II membrane protein, comprising an N-terminal cytoplasmic domain, a transmembrane domain, a putative stem region and a catalytic region containing the conserved GT1 and Gal/GalNAc-T motifs and a lectin-like domain. The *pp-GalNAc-T15* gene was localized in a draft genome sequence that was mapped to human chromosome 3 at p25. The ORF of pp-GalNAc-T15 contained at least 10 exons separated by 9 introns.

3.2. Distribution of pp-GalNAc-T15 transcripts in human tissues

The expression levels of the *pp-GalNAc-T15* transcripts in various human tissues were examined by the quantitative real-time PCR method. The *pp-GalNAc-T15* gene was expressed in most tissues examined although at different levels (Fig. 2). High levels of expression were found in small intestine, placenta, spleen, cerebral cortex, and ovary, while intermediate levels were observed in uterus, mammary gland, stomach, cerebellum, and whole brain. The expression levels were undetectable in leukocyte and very low in the other tissues tested.

3.3. Analysis of pp-GalNAc-T15 catalytic activity

Recombinant pp-GalNAc-T15 and -T2 were generated in culture media of Sf21 cells that were infected with baculoviruses carrying secretion from constructs of T15 and T2. These recombinant enzymes were purified to near homogeneity by the method described in Section 2 and were quantified by SDS–PAGE with FLAG-BAP as standard. T2 was chosen as a control for T15 because T2 sequence was well characterized and more homologous to T15 than T7. The acceptor substrate specificities of pp-GalNAc-T15 and -T2 presented in Table 1 were obtained with a variety of peptides. pp-GalNAc-T15 showed the strongest activity toward the Muc1a peptide, however, T15 showed considerably less catalytic efficiency than T2 that is arguably the most robust enzyme among pp-GalNAc-Ts.

Furthermore, detailed comparative analysis of the catalytic activities of pp-GalNAc-T15 and -T2 was performed for the Muc1a peptide, as well as the sugar donor UDP-GalNAc. The kinetic constants of GalNAc-T15 and -T2 are summarized in Table 2. Purified pp-GalNAc-T15 exhibited a K_m value of 0.016 mM for UDP-GalNAc and a K_m value of 0.910 mM for Muc1a. These values are lower than those detected for T2 (K_m of 0.081 mM for UDP-GalNAc and 0.450 mM for the Muc1a peptide substrate).

To further analyze catalytic specificity, pp-GalNAc-T15 and -T2 were incubated with Muc5AC peptide and analyzed by HPLC. Muc5AC peptide is a substrate broadly used for testing pp-GalNAc-T activities. As shown in Figs. 3A and 4A, the substrate peak (S) was gradually consumed and the product peaks (P) appeared with the extended incubation time. The

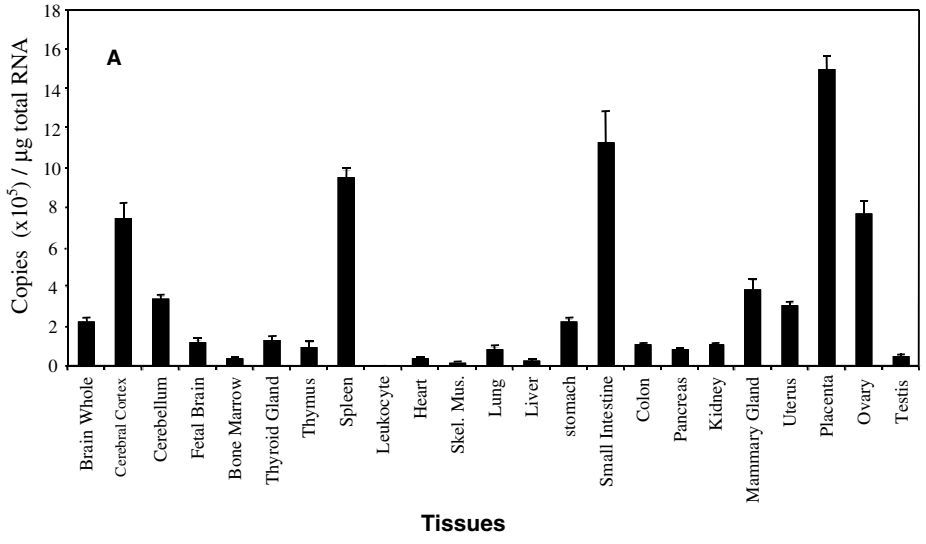

Fig. 2. Quantitative analysis of pp-GalNAc-T15 transcripts in various human tissues by real time PCR. Standard curves for pp-GalNAc-T15 and GAPDH were generated by serial dilution of each plasmid DNA. The expression level of the *pp-GalNAc-T15* transcript was normalized to that of the GAPDH transcript, which was measured in the same cDNAs. Values were expressed as copy numbers of the target gene in 1 μg of total RNA. Data were obtained from triplicate experiments and indicated as means ± SD.

Table 2
Kinetic constants of purified recombinant GalNAc-transferases

Substrate sequence	GalNAc-T15		GalNAc-T2	
	K_m (mM)	V_{max} (nmol/min/μg)	K_m (mM)	V_{max} (nmol/min/μg)
Acceptor substrate				
Mucla AHGVVTSAPDTR	0.910	0.062	0.450	0.404
Donor substrate				
UDP-GalNAc	0.016	0.087	0.081	0.0.384

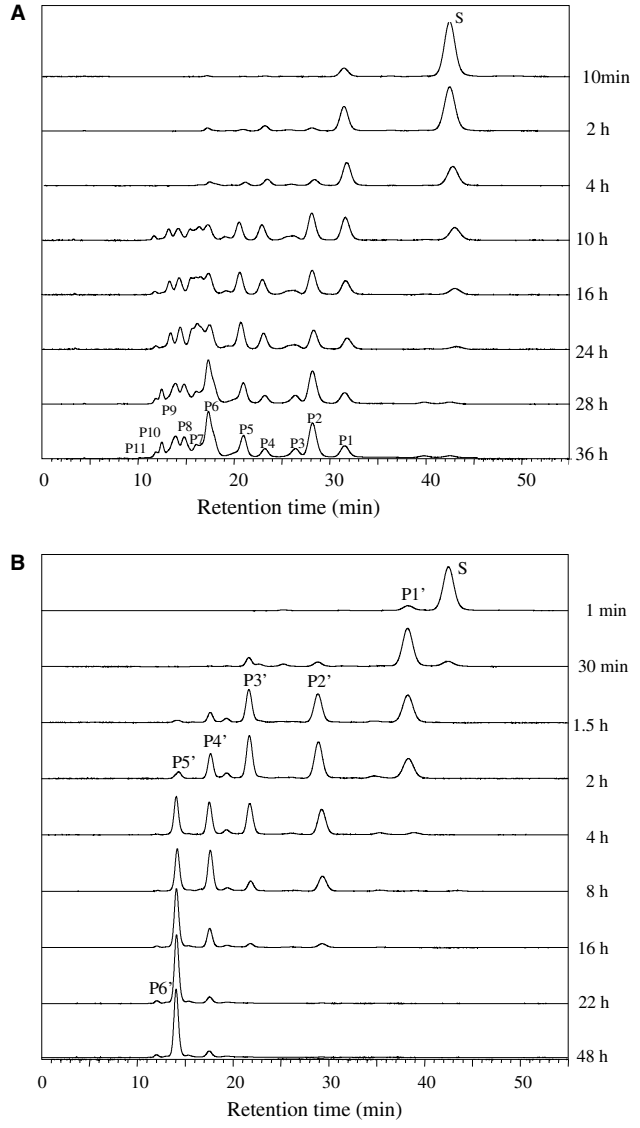

Fig. 3. HPLC analysis of the in vitro O-glycosylation of the Muc5AC peptide by pp-GalNAc-T15 and -T2. (A,B) The elution profiles of the Muc5Ac peptide glycosylated by pp-GalNAc-T15 and -T2, respectively. Peak S corresponds to the original substrate, Muc5AC peptide. Peaks P1–P11 correspond to the products generated by pp-GalNAc-T15. Peaks P1′–P6′ correspond to the products generated by pp-GalNAc-T2. The incubation time is indicated on the right of the panels.

HPLC patterns of pp-GalNAc-T15 and -T2 no longer changed even after a prolonged incubation of 24 h or the addition of fresh enzymes and UDP-GalNAc after 24 h incubation. This indicated that the peaks at 36 or 48 h are the final products of pp-GalNAc-T15 and -T2, respectively. The elution profiles of reaction products on HPLC generated by pp-GalNAc-T15 and -T2 were totally different. At least 11 product peaks, named P1–P11, were generated by pp-GalNAc-T15, and the major peaks were P2 and P6 (Fig. 3A). Whereas, pp-GalNAc-T2 generated 6 product peaks, named P1′–P6′, with P5′, the major product (Fig. 3B). The retention time of the initial product peak P1′ (38.2 min) was apparently different from that of P1 (31.5 min). These results demonstrated that pp-GalNAc-T15 and -T2 exhibited different substrate specificities.

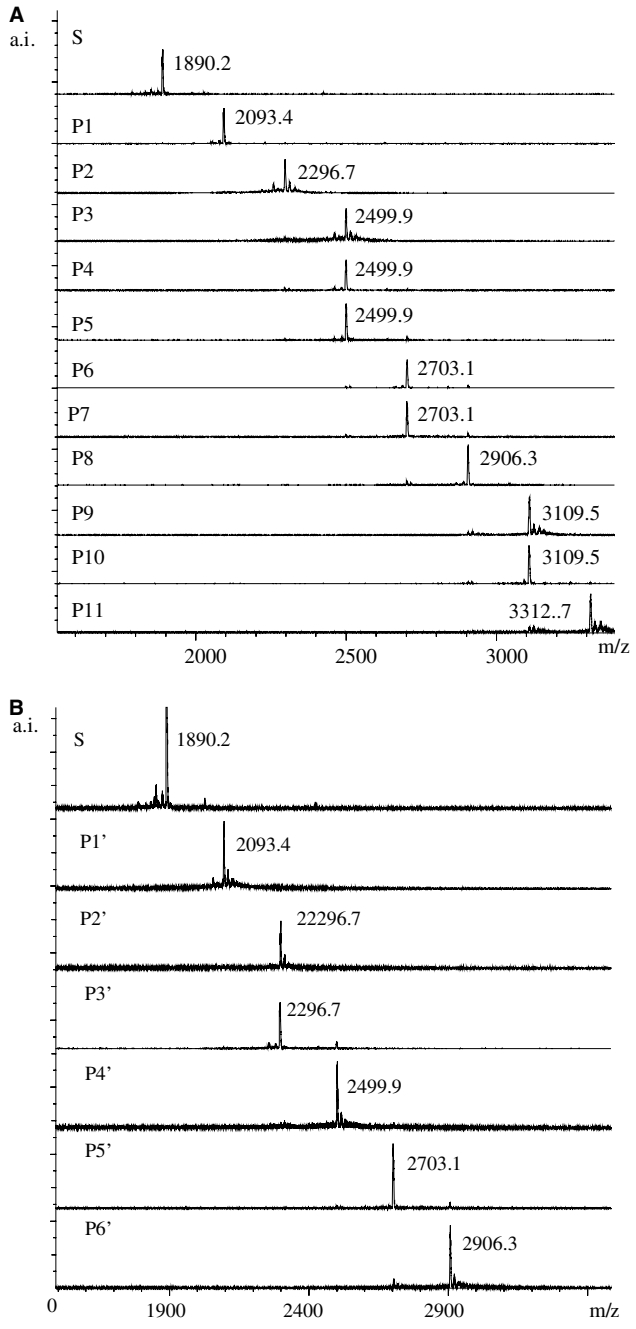

Fig. 4. MALDI-TOF-MS analysis of reaction products of pp-GalNAc-T15 (A) and -T2 (B). Each peak of the reaction products (P1–P11 and P1′–P6′) and substrate (S) in Figs. 3A and B was fractionated by HPLC and subjected to MALDI-TOF-MS analysis.

3.4. Determination of the number of GalNAc residues attached to the substrate peptide by MS

To identify the number of GalNAc residues attached to the Thr/Ser of the substrate peptide, Muc5AC, each HPLC peak of reaction products generated by GalNAc-T15 and -T2 was fractionated, and then subjected to MALDI-TOF-MS analysis. MS spectra of all peaks for pp-GalNAc-T15 and -T2 are presented in Figs. 4A and B, respectively. $[H]^-$ ion was observed as the major ion from all derivatives. From the MS analysis, the molecular weight of Muc5AC and GalNAc is 1890.2 and 203.2,

respectively. P1 had a molecular mass of 2093.4, which showed that the fraction corresponded to Muc5AC bearing one glycosylated residue. P2 had a molecular mass of 2296.7, which showed that the fraction contained two GalNAc residues. Thus, P3, P4 and P5 are identified as tri-GalNAc Muc5AC glycopeptides, and the results are interpreted as P6 and P7 (four GalNAcs), P8 and P9 (five GalNAcs), and P10 (six GalNAcs), and finally, P11 (seven GalNAcs). As shown in Fig. 4B, pp-GalNAc-T2 transferred a maximum number of five GalNAcs to the Muc5AC peptide, P1′ (one GalNAc), P2′ and P3′ (two GalNAcs), P4′ (three GalNAcs), P5′ (4GalNAcs) and finally P6′ (five GalNAcs). Thus, pp-GalNAc-T15 and -T2 differed in the maximum number of GalNAcs they incorporated into the Muc5AC peptide.

3.5. Identification of the position of GalNAc attached to Muc5AC by peptide sequencing analysis

The glycosylated peptide in each peak in Figs. 3A and B was isolated by HPLC and subjected to peptide sequencing to determine the sites of GalNAc attachment to the Muc5AC peptide. The pathways of O-glycosylation by pp-GalNAc-T15 and -T2 are schematically presented in Figs. 5A and B, respectively. As shown in Fig. 5A, the first site of incorporation by pp-GalNAc-T15 was Thr-3 (P1), then Thr-2 (P2). P3 and P5 were products of P2 generated by the addition of GalNAc to Thr-9 and Thr-12, respectively. It is clear that P2 is not the precursor of P4, which consisted of a peptide glycosylated at Thr-3, Ser-11 and Thr-13. P3 was further glycosylated at Ser-11 to produce P6 followed by Thr-10 to yield P8. P7 was from P5 obtained by the addition of GalNAc to Thr-13. P7 was then glycosylated stepwise to yield P11 according to the following order, P9 (Ser-11), P10 (Thr-10) and P11 (Thr-9). Thus, the final product generated by pp-GalNAc-T15, P11, was glycosylated at seven sites in the Muc5AC peptide, Thr-2, Thr-3, Thr-9, Thr-10, Ser-11, Thr-12 and Thr-13.

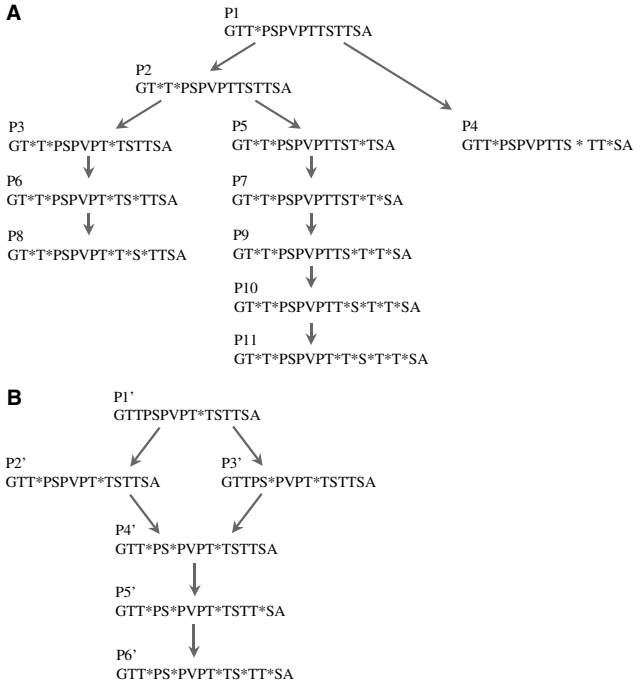

Fig. 5. Schematic representation of the predicted pathways of GalNAc incorporation into Muc5AC peptide by pp-GalNAc-T15 (A) and -T2 (B). The pathways of GalNAc incorporation into Muc5AC peptide by pp-GalNAc-T15 and -T2 were predicted based on the analysis of MS and peptide sequencing. Amino acids with asterisks are the residues to which the attachment of GalNAc was confirmed.

As shown in Fig. 5B, pp-GalNAc-T2 firstly transferred a GalNAc residue to Thr-9 (P1′), with the second site of incorporation by Thr-3 (P2′) or Ser-5 (P3′). Then, P2′ and P3′ were further glycosylated to produce P4′ by the addition of GalNAc to Ser-5 and Thr-3, respectively. P5′, having four GalNAcs at Thr-3, Ser-5, Thr-9 and Thr-13, was generated by the addition of GalNAc to Thr-13 of P4′. Finally, P6′ was formed by the addition of GalNAc to Ser-11 of P5′. Thus, the final product of pp-GalNAc-T2, P6′ had GalNAcs at five positions, Thr-3, Ser-5, Thr-9, Ser-11 and Thr-13. These results clearly indicated differences in site-specific and order-specific GalNAc incorporation by pp-GalNAc-T15 and by T2.

4. Discussion

The initiation of mucin-type O-glycosylation is catalyzed by the action of a family of pp-GalNAc-Ts. Studies over the past decade have found that pp-GalNAc-Ts display differences in both substrate specificity and tissue distribution. In the present study, we cloned and characterized a novel isoform of the GalNAc-Ts, pp-GalNAc-T15.

Mammalian pp-GalNAc-T transcripts vary in their level and distribution among human tissues. It was reported that pp-GalNAc-T1, -T2, -T4, -T7, -T8, -T10, -T12 and -T14 were expressed ubiquitously [2,5,9,10,13,15,17], whereas pp-GalNAc-T3, -T6, -T9, -T11 and -T13 had more restricted distributions [3,8,11,14,16]. By real-time PCR analysis, the pp-GalNAc-T15 transcript was broadly expressed in various tissues, although the highest level of expression was found in the small intestine, nervous system and female reproductive system. These tissues are where the *Muc5AC* gene is also highly expressed [22], suggesting that Muc5AC is a potential substrate for T15.

In addition to differential tissue distribution, distinct, although partly overlapping, substrate specificities of pp-GalNAc-Ts are considered to contribute to the O-glycosylation pattern in certain tissues and cells. In previous studies, apoproteins, peptidic fragments and synthetic peptides have been used to evaluate the O-glycosylation reactions in vitro in an attempt to define the acceptor substrate specificities for the different pp-GalNAc-Ts [26]. In the present study, we used synthetic peptides as acceptor substrates to compare the substrate specificities of purified recombinant pp-GalNAc-T15 and -T2. Both enzymes showed GalNAc-transferase activities toward a panel of peptide substrates listed in Table 1, but the initial velocities of pp-GalNAc-T15 were considerably lower than those of pp-GalNAc-T2 with all substrates examined. Even though pp-GalNAc-T15 showed the best and T2 showed the poorest activity with Muc1a, T15 showed clearly less activity with Muc1a than T2. The catalytic activity of pp-GalNAc-T2 was similar to that in previous reports [25,27]. Further detailed analysis of the activity revealed that pp-GalNAc-T15 had a lower K_m with Muc1a and UDP-GalNAc than pp-GalNAc-T2 but also had a lower V_{max}. These results demonstrated that pp-GalNAc-T15 had weaker catalytic

activity toward all acceptor substrates examined than pp-GalNAc-T2. It was reported that pp-GalNAc-T2 showed a K_m value of 2.13 mM with the Muc1a peptide [27], which was not consistent with the K_m value obtained in our experiment. We considered that the different expression and assay system might result in the difference.

The in vitro transferase analysis has shown that the sites of GalNAc incorporation in peptide substrates as well as the order of the incorporation differ markedly with each pp-GalNAc-T isoform even when the same sites are glycosylated [27–29]. In the present study, we found that the *pp-GalNAc-T15* and the *Muc5AC* genes have similar expression patterns. Muc5AC peptide (GTTPSPVPTTSTTSA), derived from human Muc5AC, has been considered an excellent substrate for evaluating the substrate specificity of GalNAc-Ts [22]. Thus, to gain further insight into the specificity of pp-GalNAc-T15, the sites and order of GalNAc incorporation into Muc5AC by pp-GalNAc-T15 and -T2 were analyzed.

As summarized in Figs. 5A and B, the sites and order of incorporation on Muc5AC were totally different between pp-GalNAc-T15 and -T2. The initial site of incorporation in the Muc5AC peptide by pp-GalNAc-T15 was Thr-3, followed by Thr-2, or Thr-13/Ser-11. The final product by T15 was glycosylated at seven sites in Muc5AC (GT*T*PSVPT*T*S*T*T*SA, GalNAc attachment sites marked by asterisks) by the order indicated in Fig. 5A. In contrast, the initial site of attachment to Muc5AC by pp-GalNAc-T2 was Thr-9, then Thr-3 or Thr-5, and the final product generated by pp-GalNAc-T2 was glycosylated at five sites (GTT*PS*VPT*TS*TT*SA). There is substantial evidence that the sequence motifs surrounding Ser/Thr residues play crucial roles in transfer of GalNAc by pp-GalNAc-Ts. The proline residues located at +3, +1 and /or −1 relative to the glycosylation site have positive effects, and the sequence motif *X*TP*X*P is considered to well serve as a signal for mucin-type O-glycosylation [30–35]. The hydroxy group of Thr-3, Ser-5, Thr-9, and Thr-13 in Muc5AC peptide matched such rules. We observed that the initial glycosylation site by pp-GalNAc-T15 and the most glycosylation sites by T2 were consistent with proline rule.

However, the site specificity of pp-GalNAc-T15 appears not to be merely governed by the peptide sequence around putative target sites. The initial GalNAc-attachment to Muc5AC peptide may also considerably influence subsequent glycosylation events through, presumably, conformation change in Muc5AC peptide. In the major glycosylation pathway by pp-GalNAc-T15, Thr-3 of the Thr-2/Thr-3 doublet was first glycosylated, and then, Thr-2 was glycosylated. After the Thr-12 was glycosylated, the vicinal Thr- or Ser- of GalNAc-attached sites was glycosylated one by one. This phenomenon is consistent with the activity of pp-GalNAc-T3 and -T4, which preferentially transfer GalNAc to the consecutive Thr residues in a synthetic peptide (PTTTPITTTTK) derived from Muc2 [36].

In contrast, pp-GalNAc-T2 could only transfer GalNAc to one of the two consecutive Thr- or Ser-residues, and the vicinal Thr- or Ser-residues of GalNAc-attached sites were not glycosylated by T2. This phenomenon resulted in that less GalNAcs were transferred onto Muc5AC peptide by T2 than T15. We presume that this critical difference between pp-GalNAc-T2 and T15 is generated by distinct acceptor recognitions by these two enzymes. The site preferences of pp-GalNAc-T2 on Muc5AC peptide were consistent with previous results using Muc2 peptide and IgA1 and hinge peptide as substrates [25,36].

Among the glycosylated products generated by pp-GalNAc-T15, P8 and P4 may be dead-end products of the reaction. They might be precursors of P10 and P9, respectively, however, no intermediates were found. P1–P4 also lacked intermediates. These intermediates, if exist, exist as small peaks that could not be fractionated, and such products were immediately further glycosylated. If P8 and/or P4 are dead-end products, it means that pp-GalNAc-T15 could not transfer GalNAcs to certain sites that are located next to GalNAc-attached residues.

We observed that pp-GalNAc-T15 transferred up to seven GalNAc residues to the Muc5AC peptide, a number significantly higher than that by pp-GalNAc-T2 (five GalNAcs) or any single established isoform. It was reported that three major isoforms of pp-GalNAc-Ts (murine T1, T2 and T3) transferred a maximum of three GalNAc residues into Muc5AC peptide [37], and only a pool of enzymes (the microsomal preparation of human gastric mucosa) cause all 6 threonine residues in the Muc5AC peptide to be fully glycosylated [38]. Thus, we presume that pp-GalNAc-T15 plays significant roles in in vivo O-glycosylation of Muc5AC and other substrates, especially in filling in vicinal Thr/Ser residues incorporation with other pp-GalNAc-Ts.

Acknowledgements: This work was performed as a part of the R&D Project of the Industrial Science and Technology Frontier Program (R&D for Establishment and Utilization of a Technical Infrastructure for Japanese Industry) supported by the New Energy and Industrial Technology Development Organization (NEDO).

References

[1] Homa, F.L., Hollander, T., Lehman, D.J., Thomsen, D.R. and Elhammer, A.P. (1993) J. Biol. Chem. 268, 12609–12616.
[2] White, T., Bennett, E.P., Takio, K., Sorensen, T., Bonding, N. and Clausen, H. (1995) J. Biol. Chem. 270, 24156–24165.
[3] Bennett, E.P., Hassan, H. and Clausen, H. (1996) J. Biol. Chem. 271, 17006–17012.
[4] Hagen, F.K., Ten Hagen, K.G., Beres, T.M., Balys, M.M., VanWuyckhuyse, B.C. and Tabak, L.A. (1997) J. Biol. Chem. 272, 13843–13848.
[5] Bennett, E.P., Hassan, H., Mandel, U., Mirgorodskaya, E., Roepstorff, P., Burchell, J., Taylor-Papadimitriou, J., Hollingsworth, M.A., Merkx, G., van Kessel, A.G., Eiberg, H., Steffensen, R. and Clausen, H. (1998) J. Biol. Chem. 273, 30472–30481.
[6] Ten Hagen, K.G., Hagen, F.K., Balys, M.M., Beres, T.M., Van Wuyckhuyse, B. and Tabak, L.A. (1998) J. Biol. Chem. 273, 27749–27754.
[7] Ten Hagen, K.G., Tetaert, D., Hagen, F.K., Richet, C., Beres, T.M., Gagnon, J., Balys, M.M., VanWuyckhuyse, B., Bedi, G.S., Degand, P. and Tabak, L.A. (1999) J. Biol. Chem. 274, 27867–27874.
[8] Bennett, E.P., Hassan, H., Mandel, U., Hollingsworth, M.A., Akisawa, N., Ikematsu, Y., Merkx, G., van Kessel, A.G., Olofsson, S. and Clausen, H. (1999) J. Biol. Chem. 274, 25362–25370.
[9] Bennett, E.P., Hassan, H., Hollingsworth, M.A. and Clausen, H. (1999) FEBS Lett. 460, 226–230.
[10] White, K.E., Lorenz, B., Evans, W.E., Meitinger, T., Strom, T.M. and Econs, M.J. (2000) Gene 246, 347–356.
[11] Toba, S., Tenno, M., Konishi, M., Mikami, T., Itoh, N. and Kurosaka, A. (2000) Biochim. Biophys. Acta 1493, 264–268.

[12] Ten Hagen, K.G., Bedi, G.S., Tetaert, D., Kingsley, P.D., Hagen, F.K., Balys, M.M., Beres, T.M., Degand, P. and Tabak, L.A. (2001) J. Biol. Chem. 276, 17395–17404.
[13] Cheng, L., Tachibana, K., Zhang, Y., Guo, J., Kahori Tachibana, K., Kameyama, A., Wang, H., Hiruma, T., Iwasaki, H., Togayachi, A., Kudo, T. and Narimatsu, H. (2002) FEBS Lett. 531, 115–121.
[14] Schwientek, T.J., Bennett, E.P., Flores, C., Thacker, J., Hollman, M., Reis, C.A., Behrens, J., Mandel, U., Keck, B., Schafer, M.A., Hazelmann, K., Zubarev, R., Roepstorff, P., Hollingsworth, M.A. and Clausen, H. (2002) J. Biol. Chem. 277, 22623–22638.
[15] Guo, J.M., Zhang, Y., Cheng, L., Iwasaki, H., Wang, H., Kubota, T., Tachibana, K. and Narimatsu, H. (2002) FEBS Lett. 524, 211–218.
[16] Zhang, Y., Iwasaki, H., Wang, H., Kudo, T., Kalka, T.B., Hennet, T., Kubota, T., Cheng, L., Inaba, N., Gotoh, M., Togayachi, A., Guo, J., Hisatomi, H., Nakajima, K., Nishihara, S., Nakamura, M., Marth, J.D. and Narimatsu, H. (2003) J. Biol. Chem. 278, 573–584.
[17] Wang, H., Tachibana, K., Zhang, Y., Iwasaki, H., Kameyama, A., Cheng, L., Guo, J., Hiruma, T., Togayachi, A., Kudo, T., Kikuchi, N. and Narimatsu, H. (2003) Biochem. Biophys. Res. Commun. 300, 738–744.
[18] Nehrke, K., Hagen, F.K. and Tabak, L.A. (1998) Glycobiology 8, 367–371.
[19] Iwai, T., Inaba, N., Naundorf, A., Zhang, Y., Gotoh, M., Iwasaki, H., Kudo, T., Togayachi, A., Ishizuka, Y., Nakanishi, H. and Narimatsu, H. (2002) J. Biol. Chem. 277, 12802–12809.
[20] Gendler, S.J., Lancaster, C.A., Taylor-Papadimitriou, J., Duhig, T., Peat, N., Burchell, J., Pemberton, L., Lalani, E.N. and Wilson, D. (1990) J. Biol. Chem. 265, 15286–15293.
[21] Gum, J.R., Byrd, J.C., Hicks, J.W., Toribara, N.W., Lamport, D.T. and Kim, Y.S. (1989) J. Biol. Chem. 264, 6480–6487.
[22] Guyonnet Duperat, V., Audie, J.P., Debailleul, V., Laine, A., Buisine, M.P., Galiegue-Zouitina, S., Pigny, P., Degand, P., Aubert, J.P. and Porchet, N. (1995) Biochem. J. 305, 211–219.
[23] Williams, S.J., Wreschner, D.H., Tran, M., Eyre, H.J., Sutherland, G.R. and McGuckin, M.A.. (2001) J. Biol. Chem. 276, 18327–18336.
[24] Albone, E.F., Hagen, F.K., VanWuyckhuyse, B.C. and Tabak, L.A. (1994) J. Biol. Chem. 269, 16845–16852.
[25] Iwasaki, H., Zhang, Y., Tachibana, K., Gotoh, M., Kikuchi, N., Kwon, Y.D., Togayachi, A., Kudo, T., Kubota, T. and Narimatsu, H.. (2003) J. Biol. Chem. 278, 5613–5621.
[26] Elhammer, A.P., Kezdy, F.J. and Kurosaka, A. (1999) Glycoconj J. 16, 171–180.
[27] Wandall, H.H., Hassan, H., Mirgorodskaya, E., Kristensen, A.K., Roepstorff, P., Bennett, E.P., Nielsen, P.A., Hollingsworth, M.A., Burchell, J., Taylor-Papadimitriou, J. and Clausen, H. (1997) J. Biol. Chem. 272, 23503–23514.
[28] Iida, S., Takeuchi, H., Hassan, H., Clausen, H. and Irimura, T. (1999) FEBS Lett. 449, 230–234.
[29] Hanisch, F.G., Muller, S., Hassan, H., Clausen, H., Zachara, N., Gooley, A.A., Paulsen, H., Alving, K. and Peter-Katalinic, J. (1999) J. Biol. Chem. 274, 9946–9954.
[30] Asada, M., Orikasa, N., Yoneda, A., Oda, Y., Ota, K. and Imamura, T. (1999) Glycoconj J. 16, 321–326.
[31] Yoshida, A., Suzuki, M., Ikenaga, H. and Takeuchi, M. (1997) J. Biol. Chem. 272, 16884–16888.
[32] Hansen, J.E., Lund, O., Engelbrecht, J., Bohr, H. and Nielsen, J.O. (1995) Biochem. J. 308, 801–813.
[33] Wilson, I.B., Gavel, Y. and von Heijne, G. (1991) Biochem. J. 275, 529–534.
[34] Hennebicq, S., Tetaert, D., Soudan, B., Boersma, A., Briand, G., Richet, C., Gagnon, J. and Degand, P. (1998) Glycoconj J. 15, 275–282.
[35] O'Connell, B.C., Hagen, F.K. and Tabak, L.A. (1992) J. Biol. Chem. 267, 25010–25018.
[36] Kato, K., Takeuchi, H., Miyahara, N., Kanoh, A., Hassan, H., Clausen, H. and Irimura, T. (2001) Biochem. Biophys. Res. Commun. 287, 110–115.
[37] Tetaert, D., Richet, C., Gagnon, J., Boersma, A. and Degand, P. (2001) Carbohydr. Res. 333, 165–171.
[38] Soudan, B., Hennebicq, S., Tetaert, D., Boersma, A., Richet, C., Demeyer, D., Briand, G. and Degand, P. (1999) J. Chromatogr. B Biomed. Sci. Appl. 729, 65–74.